21st INTERNATIONAL DEPENDENCY AND STRUCTURE MODELING CONFERENCE, DSM 2019
MONTEREY, CA, USA, 23 – 25 September, 2019

Matrix-Based Multivariate Analysis of Survey Data on Potentials for the Collaboration of Design and Simulation

Sebastian Schweigert-Recksiek, Christian Koch, Udo Lindemann

[1]Technical University of Munich, Germany

Abstract: Companies are increasingly forced to assert themselves in the market through efficient product development. Since use and potential of mechanical simulations have increased in recent years, many companies find it difficult to integrate the corresponding departments efficiently in the development process. This paper uses contingency analysis and cross tabulation to find patterns in the data of a survey-based study on the state of collaboration between design and simulation to identify significant relationships between the variables in the data set and thus describe the interdependence between the corresponding barriers and improvement measures for these departments. Using a domain mapping approach, it was possible to link suitable improvement measures to the corresponding barriers. By clustering response patterns, typical industry situations leading either to efficient or inefficient collaboration were identified. The selected methods were suitable to identify relevant connections that can help companies to choose the right measures to improve their specific situation.

Keywords: collaboration, product development process, statistical analysis, cluster analysis, domain mapping matrix

1 Introduction

In the course of digitization and increasing product complexity, the use of CAD (computer-aided design) and CAE (computer-aided engineering) systems has become indispensable in product development (Schlenkrich, 2015; Frank et al., 2007). Simulations are also being used in the early stages of development to avoid expensive tests and to shorten development cycles (Sippel, 2009; Norris, 2010). Until now, most research on the integration of simulation into the product development process has focused on technical aspects (Motte et al., 2014; Gujarathi and Ma, 2011). Yet the necessary collaboration and especially communication within and across engineering departments resulting from an increased use of simulation in development is equally critical to a successful integration (Schweigert et al., 2016). On this topic far less research has been conducted so far (Herfeld et al., 2005). This paper focuses on data of a previous study on the state of collaboration between design and simulation departments (first results were published in Schweigert et al., 2017) and tries to discover significant patterns via methods of multivariate data analysis and structural complexity management. The main goal is to derive improvement potential for the collaboration from interdependencies found in the data. Interdependencies of barriers with boundary conditions and improvement measures will be looked for. Since barriers in collaboration are the result of multiple influencing factors it is often complex to identify appropriate improvement measures (Schweigert-Recksiek and Lindeman, 2018).

By analysing the survey data, characteristic industry situations can be identified and specific measures can be linked to appearing barriers in the process of collaboration.

2 State of the Art and Research

2.1 Dimensions and State of Collaboration of Design and Simulation

Deubzer et al. (2005) have identified the four core dimensions of the holistic integration of design and simulation departments in the product development process: product, people data, and tools. The latter three aspects form the fifth dimension, process, which should be treated as a stand-alone dimension (Kreimeyer et al., 2005). Schweigert et al. (2017) conducted an industry survey on the state of collaboration of design and simulation departments within these dimensions. The design of the survey was based on a previous survey from 2006 and consisted of 31 questions in five parts: process, design, simulation, improvement opportunities, and general information. 73 usable records could be obtained from the online survey, the participants were design and simulation experts working mainly in mechanical engineering and the automotive sector. A main result of the survey was that more than half of the participants see potential to further improve collaboration as shown in Figure 1 (Question: "Do simulation and design work together efficiently in the current product development process in your opinion?").

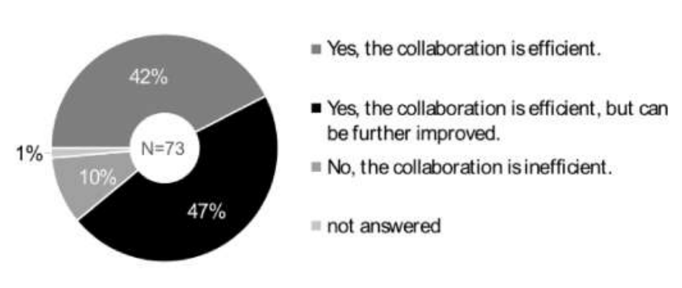

Figure 1. State of collaboration according to the survey in 2016 (Schweigert et al., 2017)

As in Maier (2007), communication in this research refers to the interaction between people and the transmission of information in a social and organizational context. It is part of collaboration, defined as the act of working together in a project or any other sort of goal-oriented activity. This is based on the "3C Collaboration Model" by Fuks et al. (2008), in which collaboration includes communication, coordination, and cooperation.

2.2 Barriers and Improvement Measures of Collaboration

Schweigert-Recksiek and Lindemann (2018) identified barriers between design and simulation departments by conducting interviews with 16 experts from 15 companies within the German engineering industry. A barrier was defined as everything that affects personal costs, computing time, or simulation results negatively, leads to unnecessary effort, or prevents necessary collaboration at all. In a workshop with collaborative researchers from social and communications sciences the connection between the empirical barriers and general barriers interdisciplinary collaboration by Eppler (2007) was built, resulting in the in a total of 20 barriers as shown in Figure 2 assigned to their respective dimensions. A full description of the methodology can be found in Schweigert-Recksiek and Lindemann (2018).

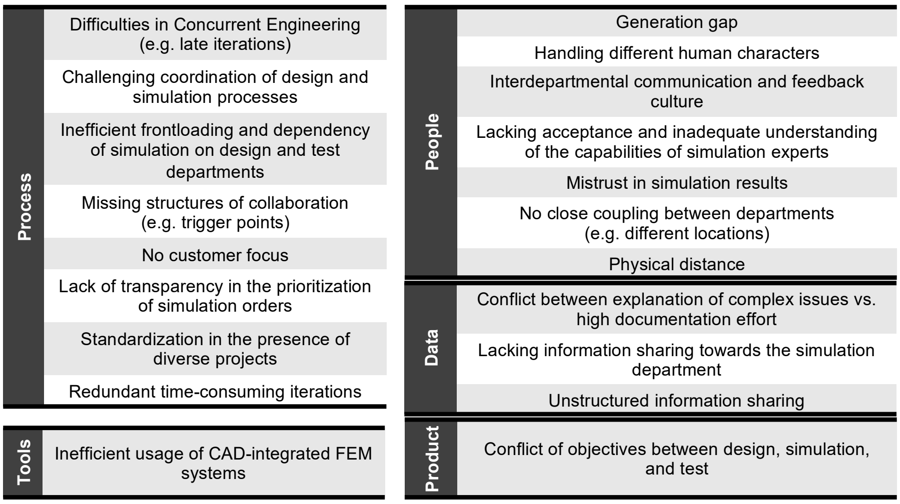

Figure 2. Barriers in collaboration between design and simulation (Schweigert-Recksiek and Lindemann, 2018)

Analogous to the barriers, 16 groups of suitable improvement measures were identified from interviews and literature review with a similar approach. They are based on a set of 120 recommendations for communication in engineering design as listed in Maier et al. (2011). An improvement measure was defined as a concrete action that can be undertaken to improve the collaboration of the design and simulation department.

3 Methodology

3.1 Research Problem und Objective

Since previous research on the integration of simulation in the product development process has focused mainly on tools and data, there is a need to further investigate the dimensions people and process (Motte et al., 2014; Schweigert et al., 2016). In particular, statistically significant statements about the factors for collaboration and especially communication between the involved departments based on empirical surveys are missing. Therefore, this empirical study aims to describe interdependencies between different factors for collaboration between design and simulation departments and to identify typical industry situations based on statistical analysis to suggest suitable improvement measures.

According to this, the following research questions guiding the analysis were derived:

RQ 1. Which statements for the improvement of the collaboration of design and simulation can be derived from the present data set?

> *RQ 1.1 Which statistical analysis methods are suitable for identifying further correlations in this dataset?*
>
> *RQ 1.2 Which of the identified correlations in the dataset can be used to further improve the collaboration between design and simulation?*

3.2 Research Design

The statistical analysis is based on the data of a previously conducted study. An overview of the scientific approach for the analysis is shown in Figure 3. As a first step, the data exported from the online-questionnaire was prepared and cleaned to obtain a dataset suitable for analysis. The selected analytical methods, the contingency and cluster analysis, were implemented in R and significant correlations and clusters were identified in the dataset. The identified patterns were linked to further research findings (cf. Schweigert-Recksiek and Lindemann, 2018) in order to interpret and evaluate the identified interdependencies and characteristic situations.

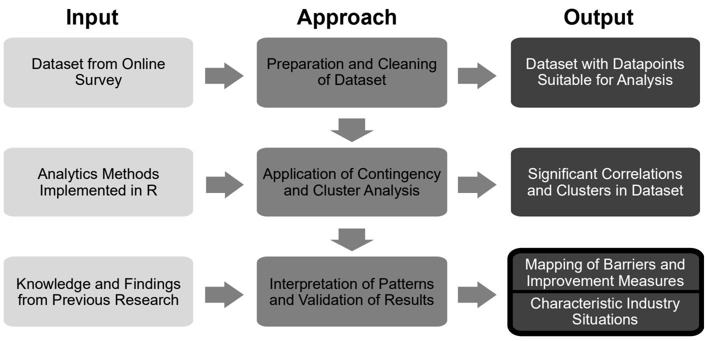

Figure 3. Scientific approach for the statistical analysis of the survey data

3.3 Statistical Methods

Due to the single-choice character of the survey, the available dataset consisted mainly of nominally scaled variables. Therefore, the contingency and cluster analysis were selected as methods for the statistical analysis.

3.3.1 Contingency Analysis and Cross Tabulation

The contingency analysis is used to evaluate the relationships between nominally scaled variables (Kuckartz et al., 2010). An important tool for contingency analysis is the cross or contingency table. It is used to examine the frequency distribution of two variables and to find interrelations between them. If the observed values deviate significantly from the expected values in the cross table, a correlation probably exists between the two variables (Backhaus et al., 2016). As a measure for the strength of the correlation, the *Cramer's V-value* of the cross table can be calculated. To evaluate the statistical significance of the correlation, the *p-value* is used. It is defined as the probability for the value to turn out equal or more extreme than the value actually observed if no correlation exists.

3.3.2 Cluster Analysis

The goal of the cluster analysis is to group the examined objects into groups (clusters) which are internally as homogeneous as possible concerning their characteristics and externally as different as possible compared to the other clusters in the dataset (Runkler, 2015). The hierarchical cluster analysis, which was selected for the analyses in this paper, consists of three steps: proximity calculation of the examined objects, performing the fusion algorithm, and the determination the optimal number of clusters (Agresti, 2012). Due to the nominal data set, the *gower*- and *chi-squared-distance* should be used as proximity measures (Backhaus et al., 2016; Gower, 1971). Concerning the fusion algorithm, the complete linkage algorithm provides the best results. The optimal number of clusters can be determined visually with the help of the cluster dendrogram and the elbow-criterion (Struyf, 1997).

4 Results

4.1 Identification of Significant Correlations

To identify significant correlations between the 23 variables, each representing the response behaviour of one survey question as listed in Schweigert et al. (2017), the strength of the correlation as well as its statistical significance were determined. As shown in Figure 4, a *Cramer's V* of 0.3 was selected as a lower limit for the strength of the correlation and a *p-value* of 0.2 as an upper limit for its significance.

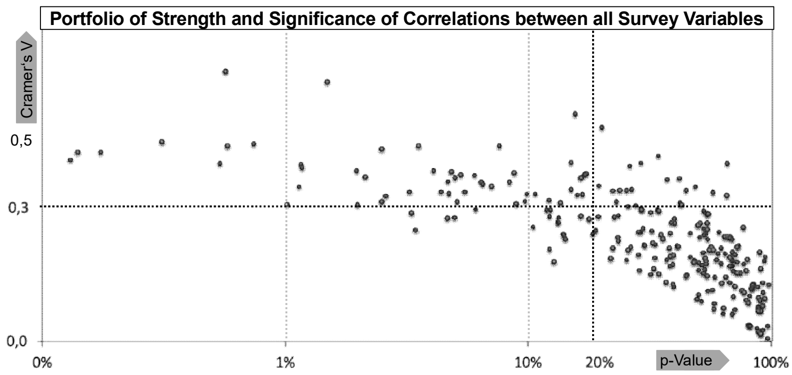

Figure 4. Identification of significant correlations via the contingency analysis

Using this approach, 40 out of the possible 253 correlations between two questions each could be identified as strong and significant for further investigation, as their values for both *Cramer's V* and *p-value* were above or below the threshold respectively.

4.2 Interdependencies of Barriers, Improvement Measures, and Boundary Conditions

To describe the interdependencies between the barriers, improvement measures, and boundary conditions, cross tabulation with the help of mosaic plots was used. Mosaic plots visualize contingency tables, displaying the occurring values as well as the deviation from the expected values graphically (Meyer et al., 2015). Exemplary the mosaic plot for the correlation of the variables "Standardized process" (Q: "Is there a standardized process of collaboration between design and simulation?") and "Efficiency of the collaboration" (Q: "Do design and simulation collaborate efficiently?") is shown in Figure 5. Three conclusions can be derived: If the standardized process is missing, collaboration will be inefficient more often than statistically expected. If the standardized process is assessed as unnecessary, collaboration will be efficient more often than statistically expected. If the standardized process is implemented, however, no distinctive effect on the efficiency will be visible. A possible interpretation of this interdependence is that a standardized process does not guarantee efficient collaboration, yet in the case of inefficient collaboration process standardization is often rated as an appropriate measure for improvement.

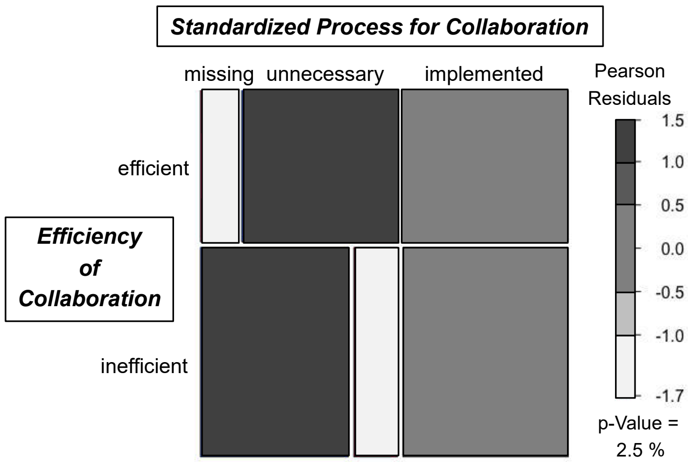

Figure 5. Mosaic plot of the correlation between efficiency of collaboration and a standardized process

4.3 Linking Improvement Measures to Barriers via Domain Mapping

To link appropriate improvement measures to barriers via the results of the contingency analysis, a matrix-based domain mapping approach can be used (cf. Maurer 2017; Eppinger and Browning, 2012).

A first matrix assigning the previously identified barriers to the corresponding variables of survey dataset is created. A second matrix contains all significant correlations between the survey variables from the contingency analysis (upper left corner in Figure 4). A third matrix assigns improvement measures to the corresponding survey variables. As shown in Figure 6, a multiplication of the three matrices results in a matrix linking improvement measures to barriers via the significant relationships identified by the contingency analysis.

The resulting connections between barriers and improvement measures was compared to connections drawn from the expert interviews in Schweigert-Recksiek and Lindemann (2018), leading to a final set of empirically based mappings.

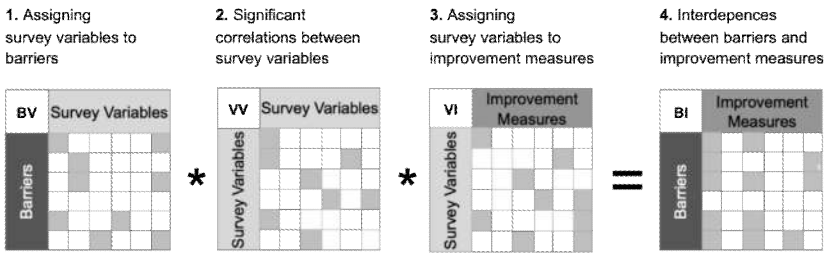

Figure 6. Matrix multiplication to link barriers and improvement measures

4.4 Identification of Characteristic Industry Situations

By clustering the response patterns of the survey participants, characteristic situations in the industry concerning the variables in the dataset are identifiable. Several clustering approaches were evaluated. The best results were obtained using the *gower-* and *chi-squared-distance* measures (exemplary dendrogram with optimal number of 5 clusters shown in Figure 7) when combining them with agglomerative or divisive clustering algorithms and a complete-linkage fusion algorithm. Since clustering using all possible variables produced only hardly comparable clusters, clustering was also conducted using two smaller sets of variables, which were identified as strongly linked to the efficiency of the collaboration beforehand via cross tabulation.

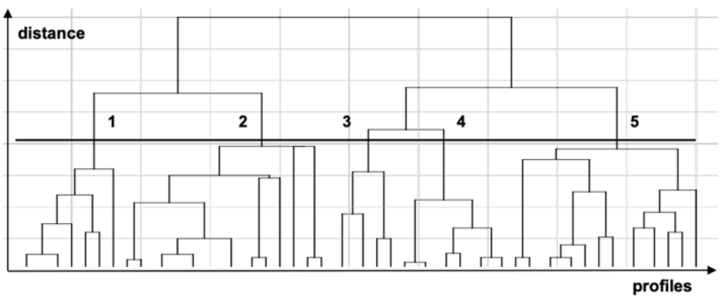

Figure 7. Dendrogram of agglomerative clustering of the dataset with Gower-Distance

All clusters were subsequently analysed and compared concerning their characteristics. Only clusters with a sufficient amount of homogenous properties and a clearly determinable efficiency of the collaboration were considered for further analysis. Out of the 48 obtained clusters a total of six unique clusters describing efficient collaboration and six unique clusters describing inefficient collaboration could be condensed. The clusters were further analysed and summarized by reducing them to their core statements as shown in Figure 8. For example, “Tools” in this case refers to either a common PDM/PLM system, CAD-integrated FEM systems, or preparations for simulations in CAD.

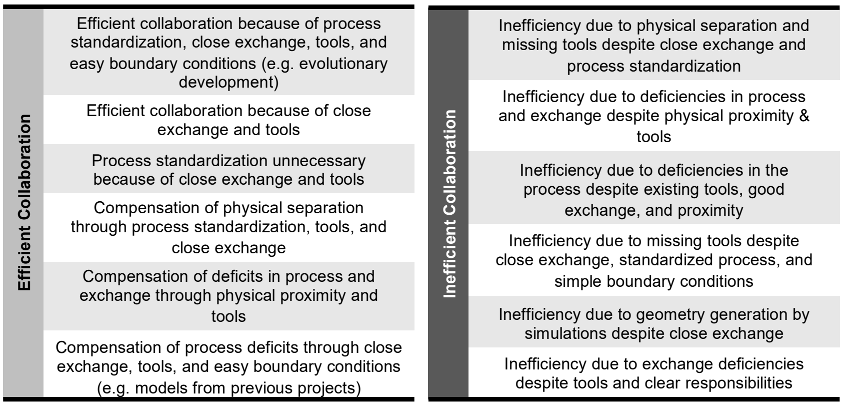

Figure 8. Characteristic industry situations identified from clusters

5 Conclusion

5.1 Discussion

Despite the meaningful results presented in the previous section, some limitations remain. A main issue conducting the statistical analysis was the sample size given by the number of survey participants. The number of 73 usable datasets is comparatively small for most methods of data exploration and because of incomplete response patterns the sample size for some variables was even smaller. This results in an increased probability of statistical errors: non-existent correlations might accidentally be detected, or existent ones might be overlooked. However, considering that the study is an expert survey, the number of participants is comparatively large. Another possible issue was that the responding behaviour of the participants cannot always be assumed to be perfectly objective. The participants are experts who do not judge their company from the outside but are part of the process and therefore give their subjective assessment. However, because of their expertise and state as experts their opinions are amongst the most trustworthy available and must thus be relied upon. In addition, the assignment of survey variables to barriers and improvement measures was not always possible unambiguously. For example, non-existent improvement measures can also be interpreted as barriers for collaboration. This results in a certain degree of fuzziness in the connected analytical approaches.

Despite the aforementioned problems due to characteristics of the survey, the dataset could nevertheless be analysed sufficiently well. The selected methods of contingency and cluster analysis are each tailored to the handling of binary and nominal data points. It was thus possible to identify clear patterns and correlations both by examining the relationships of the variables and by clustering the response patterns.

5.2 Outlook

Further analyses of the available dataset can include other methods of categorical data analysis like logistic regression or correspondence analysis to uncover other hidden patterns. Next steps in research include further expert interviews to validate the statistical findings as well as to generate further data to test if similar patterns occur.

The results in this paper are part of a framework presented in Schweigert et al. (2017) that generates system graphs on the collaboration of design and simulation departments. The mapping of barriers and suitable improvement measures makes it possible to semi-automatically find improvement opportunities after recording the state of collaboration in a company, resulting in appropriate measures to improve the collaboration.

References

Agresti, A., 2012**.** Categorical Data Analysis. Wiley & Sons, New York.

Backhaus, K., Erichson, B., Plinke, W. and Weiber, R., 2016. Multivariate Analysemethoden - Eine anwendungsorientierte Einführung. Springer, Berlin.

Deubzer, F., Herfeld, U., Kreimeyer, M. and Lindemann, U., 2005. A structured holistic approach for the integration of CAD and CAE environments. ProSTEP iViP Science Days 2005, pp. 1–10.

Eppinger, S., Browning, T., 2012. Design Structure Matrix Methods and Applications. MIT Press, Cambridge.

Gujarathi, G., Ma, Y., 2011. Parametric CAD/CAE integration using a common data model. Journal of Manufacturing Systems, Vol. 30, No. 3, 2011, pp. 118–132.

Gower, J., 1971. A general coefficient of similarity and some of its properties. Biometrics, Vol. 27, Nr. 4., pp. 857-871.

Eppler, M., 2007. Knowledge communication problems between experts and decision makers: An overview and classification. The Electronic Journal of Knowledge Management, 53(3), pp. 291-300.

Eriksson, M., Petersson, H., Bjärnemo, R. and Motte, D., 2014. Interaction between computer-based design analysis activities and the engineering design process. The 13th International Design Conference. The Design Society, Glasgow, pp. 1283–1296.

Frank, T., Herfeld, U., Kreimeyer, M. and Lindemann, U., 2007. Parametrische Optimierung im CAD-CAE-Prozess der Karosserieentwicklung. ANSYS Conference and 25th CADFEM Users Meeting.

Fuks, H., Raposo, A., Gerosa, M. A., Pimentel, M. and Lucena, C., 2008. The 3C Collaboration Model, In: Kock, N. Ed., Encyclopaedia of e-Collaboration, Information Science Reference, Hershey, pp. 637-644.

Herfeld, U., Deubzer, F., Kreimeyer, M. and Lindemann, U., 2005. Application of CAD-Integrated FEM Analysis as a Step towards Better Integration of Simulation into the Design Process. The 9th International Research/Expert Conference Trends in the Development of Machinery and Associated Technology.

Kuckartz, U., Rädiker, S., Ebert, T., and Schehl, J., 2010. Statistik - Eine verständliche Einführung. VS Verlag, Wiesbaden.

Kreimeyer, M., Deubzer, F., Herfeld, U. and Lindemann, U., 2005. Holistic integration of CAD and CAE: Analysis and combination of diverse current approaches. The 9th International Research/Expert Conference Trends in the Development of Machinery and Associated Technology, pp. 1081–1084.

Maier, A. M., 2007. A grid-based assessment method of communication in engineering design. Cambridge, University of Cambridge.

Maurer, M., 2017. Complexity Management in Engineering Design. Springer, Berlin.

Meyer, D., Zeileis, A., Hornik, K., 2015. The Strucplot Framework: Visualizing Multi-way Contingency Tables with vcd. Journal of Statistical Software, 2006(3), pp. 1-48.

Motte, D., Eriksson, M., Petersson, H., and Bjärnemo, R., 2014. Integration of the computer-based design analysis activity in the engineering design process - A literature survey. Proceedings of the 10th International Symposium on Tools and Methods of Competitive Engineering. Delft University of Technology, pp. 1181-1194.

Norris, M., 2010. The Value of Simulation Process & Data Management - Lessons from a Decade of Production Experience. NAFEMS 2010 European Conference: Simulation Process and Data Management.

Runkler, T. A., 2015. Data Mining - Modelle und Algorithmen intelligenter Datenanalyse . Springer Vieweg. Wiesbaden.

Petersson, H., Motte, D., and Bjärnemo, R., 2015. Using Templates to Support the Engineering Designer Performing Computer-Based Design Analysis. International Mechanical Engineering Congress and Exposition, Vol. 11, pp. V011T14A002.

Reicheneder, J., 2015. 15 Years SPDM at BMW. NAFEMS 2015 European Simulation Process and Data Management Conference.

Schweigert, S., d'Albert, H., and Lindemann, U., 2016. Enhancement of Collaboration and Communication between Design and Simulation Departments by Methods of Requirements Engineering. The 14th International Design Conference. The Design Society, Glasgow, pp. 1397-1406.

Schweigert, S., Xia, M., and Lindemann, U., 2017. An empirical survey on efficiency improvement for the collaboration between design and simulation departments, The 21st International Conference on Engineering Design. The Design Society, Glasgow, pp. 337-346.

Schweigert-Recksiek, S., Lindemann, U., 2018. Improvement Opportunities for the Collaboration of Design and Simulation Departments - An Interview Study. The international Design Conference on Engineering Design. The Design Society, Zagreb, pp. 905-916

Schlenkrich, M., 2015. Handling the Constant Evolution of Simulation with SPDM Solutions. NAFEMS 2015 European Simulation Process and Data Management Conference.

Sippel, H., 2009. CAE at Different Stages in Product Development. NAFEMS 2009 European Simulation Process and Data Management Conference.

Struyf, A., Hubert, M., Roussew, P., 1997. Clustering in an Object-Oriented Environment. Journal of Statistical Software, 1(4), pp. 1-30.

Acknowledgements

The authors thank the survey participants and interview partners for the provided insights. The authors also thank Minghai Xia for his part in designing and conducting the survey, Felix

Wiedenhaupt for his support in performing further interviews, Elena Muro for her insights from social and communication sciences and Michael Ponnath for the first part of the statistical analysis.

Contact: Sebastian Schweigert-Recksiek, Technical University of Munich, Department of Mechanical Engineering, Product Development and Lightweight Design, Boltzmannstrasse 15, 85748 Garching, Germany, Phone: +49-89-289-15126, Fax: +49-89-289-15144, schweigert.recksiek@tum.de, www.mw.tum.de/lpl